

Min- and Max- Relative Entropies and a New Entanglement Monotone

Nilanjana Datta*

Statistical Laboratory, DPMMS, University of Cambridge, Cambridge CB3 0WB, UK

(Dated: November 25, 2009)

Two new relative entropy quantities, called the *min- and max-relative entropies*, are introduced and their properties are investigated. The well-known min- and max- entropies, introduced by Renner [1], are obtained from these. We define a new entanglement monotone, which we refer to as the *max-relative entropy of entanglement*, and which is an upper bound to the relative entropy of entanglement. We also generalize the min- and max-relative entropies to obtain *smooth min- and max- relative entropies*. These act as parent quantities for the *smooth Rényi entropies* [1], and allow us to define the analogues of the mutual information, in the Smooth Rényi Entropy framework. Further, the spectral divergence rates of the Information Spectrum approach are shown to be obtained from the smooth min- and max-relative entropies in the asymptotic limit.

PACS numbers: 03.65.Ud, 03.67.Hk, 89.70.+c

Keywords: quantum relative entropy, smooth Rényi entropies, spectral divergence rates, information spectrum, entanglement monotone

I. INTRODUCTION

One of the fundamental quantities in Quantum Information Theory is the relative entropy between two states. Other entropic quantities, such as the von Neumann entropy of a state, the conditional entropy and the mutual information for a bipartite state, are obtainable from the relative entropy. Many basic properties of these entropic quantities can be derived from those of the relative entropy. The strong subadditivity of the von Neumann entropy, which is one of the most powerful results in Quantum Information Theory, follows easily from the monotonicity of the relative entropy. Other than acting as a parent quantity for other entropic quantities, the relative entropy itself has an operational meaning. It serves as a measure of distinguishability between different states.

The notion of relative entropy was introduced in 1951, in Mathematical Statistics, by Kullback and Leibler [2], as a means of comparing two different probability distributions. Its extension to the quantum setting was due to Umegaki [3]. The classical relative entropy plays a role similar to its quantum counterpart. Classical entropic quantities such as the Shannon entropy of a random variable, the conditional entropy, the mutual information and the joint entropy of a pair of random variables are all obtainable from it.

More recently, the concept of relative entropy has been generalized to sequences of states, in the so-called Information Spectrum Approach [4, 5, 6, 7, 8, 9]. The latter is a powerful method which enables us to evaluate the optimal rates of various information theory protocols, without making any assumption on the structure of the sources, channels or (in the quantum case) the en-

tanglement resources involved. In particular, it allows us to eliminate the frequently-used, but often unjustified, assumption that sources, channels and entanglement resources are memoryless. The quantities arising from the generalizations of the relative entropy in this approach are referred to as *spectral divergence rates* [see Section VII for their definitions and properties]. Like the relative entropy, they yield quantities which can be viewed as generalizations of entropy rates for sequences of states (or probability distributions, in the classical case). These quantities have been proved to be of important operational significance in Classical and Quantum Information Theory, as the optimal rates of protocols such as data compression, dense coding, entanglement concentration and dilution, transmission of classical information through a quantum channel and in the context of hypothesis testing [see e.g. [7, 8, 10, 11, 12, 13, 14, 15]]. Hence, spectral divergence rates can be viewed as the basic tools of a unifying mathematical framework for studying information theoretical protocols.

A simultaneous but independent approach, developed to overcome the limitation of the memoryless criterion is the so-called *Smooth Entropy framework*, developed by Renner et al. (see e.g. [1], [16], [17], [18], [19]). This approach introduced new entropy measures called *smooth Rényi entropies* or *smooth min- and max- entropies*. In contrast to the spectral entropy rates, the (unconditional and conditional) smooth min- and max- entropies are defined for individual states (or probability distributions) rather than sequences of states. They are non-asymptotic in nature but depend on an additional parameter ϵ , the smoothness parameter. Similar to the spectral entropies, the min- and max- entropies have various interesting properties e.g. chain rule inequalities and strong subadditivity. They are also of operational significance and have proved useful in the context of randomness extraction and cryptography.

Recently it was shown [20] that the two approaches

*Electronic address: N.Datta@statslab.cam.ac.uk

discussed above, are related in the sense that the spectral entropy rates are obtained as asymptotic limits of the corresponding smooth min- and max- entropies.

In this paper we introduce two new relative entropy quantities, namely the *min- and max- relative entropies* (and their smoothed versions), which act as parent quantities for the unconditional and conditional min- and max- entropies of Renner [1]. These new relative entropy quantities are seen to satisfy several interesting properties. Their relations to the quantum relative entropy [21] and to the quantum Chernoff bound [22] are discussed. They also allow us to define analogues of the mutual information in the *Smooth Rényi Entropy Framework*. The operational significance of the latter will be elaborated in a forthcoming paper. We define a new entanglement monotone, which we refer to as the *max-relative entropy of entanglement*, and which is an upper bound to the relative entropy of entanglement. Moreover, the smooth min- and max- relative entropies and the analogous quantities in the Quantum Information Spectrum framework, namely the spectral divergence rates, are proved to be related in the asymptotic limit. The proofs are entirely self-contained, relying only on the definitions of the entropic quantities involved, and the lemmas stated in Section II.

The min- and max- relative entropies both have interesting operational significances. The operational meaning of the min- relative entropy is given in state discrimination as the negative logarithm of the optimal error probability of the second kind, when the error probability of the first kind is required to be zero. This is explained in the proof of Lemma 12. The max- relative entropy, on the other hand is related to the optimal Bayesian error probability, in determining which one of a finite number (say M) of known states a given quantum system is prepared in. Suppose the quantum system is prepared in the k^{th} state, ρ_k , with apriori probability p_k , and the optimisation is over all possible choices of POVMs which could be made on the system to determine its state. Then the optimal Bayesian probability of error is given by

$$P_{av} = 1 - \inf_{\sigma} \max_{1 \leq k \leq M} p_k 2^{D_{max}(\rho_k || \sigma)},$$

where the infimum is taken over all possible quantum states, σ , in the Hilbert space of the system. This operational interpretation, was first provided in [23] though in a somewhat different formalism. It is also explained in [24].

We start with some mathematical preliminaries in Section II. We define (*non-smooth*) min- and max- relative entropies in Section III, and state how the unconditional and conditional min- and max- entropies are obtained from them. We also define the min- and max- mutual informations. In Section IV we investigate the properties of the new relative entropy quantities. A new entanglement monotone is introduced in Theorem 1 of Section V, and some of its properties are discussed. Next, we define the smoothed versions of the min- and max- relative

entropies in Section VI. After briefly recalling the definitions and basic properties of the spectral divergence rates in Section VII, we go on to prove the relations between them and the smooth min- and max- relative entropies in Section VIII. These are stated as Theorem 2 and Theorem 3, which along with the new entanglement monotone (Theorem 1), and the properties of the min- and max- relative entropies, constitute the main results of this paper. Our results apply to the quantum setting and thus include the classical setting as a special case.

II. MATHEMATICAL PRELIMINARIES

Let $\mathcal{B}(\mathcal{H})$ denote the algebra of linear operators acting on a finite-dimensional Hilbert space \mathcal{H} . The von Neumann entropy of a state ρ , i.e., a positive operator of unit trace in $\mathcal{B}(\mathcal{H})$, is given by $S(\rho) = -\text{Tr} \rho \log \rho$. Throughout this paper, we take the logarithm to base 2 and all Hilbert spaces considered are finite-dimensional. We denote the identity operator in $\mathcal{B}(\mathcal{H})$ by I .

In this paper we make extensive use of spectral projections. Any self-adjoint operator A acting on a finite-dimensional Hilbert space may be written in its spectral decomposition $A = \sum_i \lambda_i P_i$, where P_i denotes the orthogonal projector onto the eigenspace of A spanned by eigenvectors corresponding to the eigenvalue λ_i . We define the positive spectral projection on A as $\{A \geq 0\} := \sum_{\lambda_i \geq 0} P_i$, the projector onto the eigenspace of A corresponding to positive eigenvalues. Corresponding definitions apply for the other spectral projections $\{A < 0\}$, $\{A > 0\}$ and $\{A \leq 0\}$. For two operators A and B , we can then define $\{A \geq B\}$ as $\{A - B \geq 0\}$. The following key lemmas are useful. For a proof of Lemma 1 and Lemma 3, see [4, 7, 8]. Lemma 2 is proved in [20].

Lemma 1 *For self-adjoint operators A , B and any positive operator $0 \leq P \leq I$, we have*

$$\text{Tr}[P(A - B)] \leq \text{Tr}[\{A \geq B\}(A - B)] \quad (1)$$

$$\text{Tr}[P(A - B)] \geq \text{Tr}[\{A \leq B\}(A - B)]. \quad (2)$$

Identical conditions hold for strict inequalities in the spectral projections $\{A < B\}$ and $\{A > B\}$.

Lemma 2 *Given a state ρ_n and a self-adjoint operator ω_n , for any real γ , we have*

$$\text{Tr}[\{\rho_n \geq 2^{n\gamma} \omega_n\} \omega_n] \leq 2^{-n\gamma}.$$

Lemma 3 *For self-adjoint operators A and B , and any completely positive trace-preserving (CPTP) map \mathcal{T} , the inequality*

$$\text{Tr}[\{\mathcal{T}(A) \geq \mathcal{T}(B)\} \mathcal{T}(A - B)] \leq \text{Tr}[\{A \geq B\}(A - B)] \quad (3)$$

holds.

The trace distance between two operators A and B is given by

$$\|A - B\|_1 := \text{Tr}[\{A \geq B\}(A - B)] - \text{Tr}[\{A < B\}(A - B)] \quad (4)$$

The fidelity of states ρ and ρ' is defined to be

$$F(\rho, \rho') := \text{Tr} \sqrt{\rho^{\frac{1}{2}} \rho' \rho^{\frac{1}{2}}}.$$

The trace distance between two states ρ and ρ' is related to the fidelity $F(\rho, \rho')$ as follows (see (9.110) of [25]):

$$\frac{1}{2} \|\rho - \rho'\|_1 \leq \sqrt{1 - F(\rho, \rho')^2} \leq \sqrt{2(1 - F(\rho, \rho'))}. \quad (5)$$

We use the following simple corollary of Lemma 1:

Corollary 1 *For self-adjoint operators A , B and any positive operator $0 \leq P \leq I$, the inequality*

$$\|A - B\|_1 \leq \varepsilon,$$

for any $\varepsilon > 0$, implies that

$$\text{Tr}[P(A - B)] \leq \varepsilon.$$

We also use the “gentle measurement” lemma [26, 27]:

Lemma 4 *For a state ρ and operator $0 \leq \Lambda \leq I$, if $\text{Tr}(\rho\Lambda) \geq 1 - \delta$, then*

$$\|\rho - \sqrt{\Lambda}\rho\sqrt{\Lambda}\|_1 \leq 2\sqrt{\delta}.$$

The same holds if ρ is only a subnormalized density operator.

III. DEFINITIONS OF MIN- AND MAX-RELATIVE ENTROPIES

Definition 1 *The max- relative entropy of two operators ρ and σ , such that $\rho \geq 0$, $\text{Tr}\rho \leq 1$ and $\sigma \geq 0$, is defined by*

$$D_{\max}(\rho||\sigma) := \log \min\{\lambda : \rho \leq \lambda\sigma\} \quad (6)$$

Note that $D_{\max}(\rho||\sigma)$ is well-defined if $\text{supp } \rho \subseteq \text{supp } \sigma$. For ρ and σ satisfying $\text{supp } \rho \subseteq \text{supp } \sigma$, $D_{\max}(\rho||\sigma)$ is equivalently given by

$$D_{\max}(\rho||\sigma) := \log \mu_{\max}(\sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}}), \quad (7)$$

where the notation $\mu_{\max}(A)$ is used to denote the maximum eigenvalue of the operator A , and the inverses are generalized inverses defined as follows: A^{-1} is a generalized inverse of A if $AA^{-1} = A^{-1}A = P_A = P_{A^{-1}}$, where $P_A, P_{A^{-1}}$ denote the projectors onto the supports of A and A^{-1} respectively.

Another equivalent definition of $D_{\max}(\rho||\sigma)$ is:

$$D_{\max}(\rho||\sigma) := \log \min\{\lambda : \text{Tr}[P_+^\lambda(\rho - \lambda\sigma)] = 0\}, \quad (8)$$

where $P_+^\lambda := \{\rho \geq \lambda\sigma\}$.

Definition 2 *The min- relative entropy of two operators ρ and σ , such that $\rho \geq 0$, $\text{Tr}\rho \leq 1$ and $\sigma \geq 0$, is defined by*

$$D_{\min}(\rho||\sigma) := -\log \text{Tr}(\pi_\rho \sigma), \quad (9)$$

where π_ρ denotes the projector onto $\text{supp } \rho$, the support of ρ . It is well-defined if $\text{supp } \rho$ has non-zero intersection with $\text{supp } \sigma$.

Note that

$$D_{\min}(\rho||\sigma) = \lim_{\alpha \rightarrow 0^+} S_\alpha(\rho||\sigma), \quad (10)$$

where $S_\alpha(\rho||\sigma)$ denotes the *quantum relative Rényi entropy* of order α , with $0 < \alpha < 1$, defined by (see e.g. [21, 28]):

$$S_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1} \log \text{Tr} \rho^\alpha \sigma^{1-\alpha}. \quad (11)$$

Various properties of $D_{\min}(\rho||\sigma)$ and $D_{\max}(\rho||\sigma)$ are discussed in Section IV.

The min- and max- (unconditional and conditional) entropies, introduced by Renner in [1] can be obtained from $D_{\min}(\rho||\sigma)$ and $D_{\max}(\rho||\sigma)$ by making suitable substitutions for the positive operator σ . In particular, for $\sigma = I$, we obtain the min- and max- entropies of a state ρ , which are simply the Rényi entropies of order infinity and zero, respectively:

$$H_{\min}(\rho) = -D_{\max}(\rho||I) = -\log \|\rho\|_\infty \quad (12)$$

$$H_{\max}(\rho) = -D_{\min}(\rho||I) = \log \text{rank}(\rho). \quad (13)$$

The min- and max-entropies of a bipartite state, ρ_{AB} , relative to a state σ_B , are similarly obtained by setting $\sigma = I_A \otimes \sigma_B$:

$$\begin{aligned} H_{\min}(\rho_{AB}|\sigma_B) &:= -\log \min\{\lambda : \rho_{AB} \leq \lambda \cdot I_A \otimes \rho_B\} \\ &= -D_{\max}(\rho_{AB}||I_A \otimes \sigma_B) \end{aligned} \quad (14)$$

and

$$\begin{aligned} H_{\max}(\rho_{AB}|\sigma_B) &:= \log \text{Tr}(\pi_{AB}(I_A \otimes \sigma_B)) , \\ &= -D_{\min}(\rho_{AB}||I_A \otimes \sigma_B) \end{aligned} \quad (15)$$

In the above, π_{AB} denotes the projector onto the support of ρ_{AB} .

In addition, by considering $\sigma = \rho_A \otimes \rho_B$, we obtain the following analogues of the quantum mutual information of a bipartite state $\rho = \rho_{AB}$:

Definition 3 *For a bipartite state ρ_{AB} , the min- and max- mutual informations are defined by*

$$\begin{aligned} D_{\min}(A : B) &:= D_{\min}(\rho_{AB}||\rho_A \otimes \rho_B) \\ D_{\max}(A : B) &:= D_{\max}(\rho_{AB}||\rho_A \otimes \rho_B) \end{aligned} \quad (16)$$

Smooth versions of the min- and max- relative entropies are defined in Section VI. These in turn yield the (unconditional and conditional) smooth min- and max-entropies [1, 20] and mutual informations, upon similar substitutions for the operator σ . It was proved in [20] that the smooth min- and max- entropies are related to the spectral entropy rates used in the Quantum Information Spectrum approach [see Section VII or [4]], in the sense that the spectral entropy rates are the asymptotic limits of the smooth entropies. As discussed in Section VII, the spectral entropy rates are obtainable from two quantities, namely the inf- and sup- spectral divergence rates. In Section VIII, we prove that these spectral divergence rates are indeed asymptotic limits of the smooth min- and max- relative entropies.

IV. PROPERTIES OF MIN- AND MAX-RELATIVE ENTROPIES

The min- and max- relative entropies satisfy the following properties:

Lemma 5 *For a state ρ and a positive operator σ*

$$D_{\min}(\rho||\sigma) \leq D_{\max}(\rho||\sigma) \quad (17)$$

Proof (This is exactly analogous to the proof of Lemma 3.1.5 in [1]). Let π_ρ denote the projector onto the support of ρ , and let $\lambda \geq 0$ such that $D_{\max}(\rho||\sigma) = \log \lambda$, i.e., $\lambda\sigma - \rho \geq 0$. Then, using the fact that for positive semi-definite operators A and B , $\text{Tr}(AB) \geq 0$, we get

$$0 \leq \text{Tr}((\lambda\sigma - \rho)\pi_\rho) = \lambda\text{Tr}(\pi_\rho\sigma) - 1.$$

Hence,

$$D_{\min}(\rho||\sigma) := -\log \text{Tr}(\pi_\rho\sigma) \leq \log \lambda = D_{\max}(\rho||\sigma)$$

Lemma 6 *The min- and max- relative entropies are non-negative when both ρ and σ are states. They are both equal to zero when ρ and σ are identical states. Moreover, $D_{\min}(\rho||\sigma) = 0$ when ρ and σ have identical supports.*

Proof Due to Lemma 5, it suffices to prove that $D_{\min}(\rho||\sigma) \geq 0$, when ρ and σ are states. Note that $\text{Tr}(\pi_\rho\sigma) \leq \text{Tr}\sigma = 1$, where π_ρ denotes the projector onto the support of ρ . Hence,

$$D_{\min}(\rho||\sigma) := -\log \text{Tr}(\pi_\rho\sigma) \geq 0.$$

The rest of the lemma follows directly from the definitions (7) and (9) of the max- and min- relative entropies, respectively. ■

Lemma 7 *The min- and max- relative entropies are monotonic under CPTP maps, i.e., for a state ρ , a positive operator σ , and a CPTP map \mathcal{T} :*

$$D_{\min}(\mathcal{T}(\rho)||\mathcal{T}(\sigma)) \leq D_{\min}(\rho||\sigma) \quad (18)$$

and

$$D_{\max}(\mathcal{T}(\rho)||\mathcal{T}(\sigma)) \leq D_{\max}(\rho||\sigma) \quad (19)$$

Proof The monotonicity (18) follows directly from the monotonicity of the quantum relative Rényi entropy. For $0 < \alpha < 1$, we have [28]:

$$S_\alpha(\rho||\sigma) \leq S_\alpha(\mathcal{T}(\rho)||\mathcal{T}(\sigma)).$$

Taking the limit $\alpha \rightarrow 0^+$ on both sides of this inequality and using (10), yields (18).

The proof of (19) is analogous to Lemma 3.1.12 of [1]. Let $\lambda \geq 0$ such that $\log \lambda = D_{\max}(\rho||\sigma)$ and hence $(\lambda\sigma - \rho) \geq 0$. Since \mathcal{T} is a CPTP map, $\mathcal{T}(\lambda\sigma - \rho) = \lambda\mathcal{T}(\sigma) - \mathcal{T}(\rho) \geq 0$. Hence,

$$\begin{aligned} D_{\max}(\mathcal{T}(\rho)||\mathcal{T}(\sigma)) &:= \log \min\{\lambda' : \mathcal{T}(\rho) \leq \lambda'\mathcal{T}(\sigma)\} \\ &\leq \log \lambda = D_{\max}(\rho||\sigma). \end{aligned} \quad (20)$$

■

Lemma 8 *The min-relative entropy is jointly convex in its arguments.*

Proof The proof follows from the monotonicity of the min-relative entropy under CPTP maps (Lemma 7). Following [28], let ρ_1, \dots, ρ_n be states acting on a Hilbert space \mathcal{H} , let $\sigma_1, \dots, \sigma_n$ be positive operators in $\mathcal{B}(\mathcal{H})$ such that $\text{supp } \rho_i \subseteq \text{supp } \sigma_i$, for $i = 1, \dots, n$, and let $\{p_i\}_{i=1}^n$ denote a probability distribution. Let $\rho, \sigma \in \mathcal{B}(\mathcal{H})$.

Consider the following operators in $\mathcal{B}(\mathcal{H} \otimes \mathbb{C}^n)$

$$\begin{aligned} A &:= \sum_{i=1}^n p_i A_i = \sum_{i=1}^n p_i |i\rangle\langle i| \otimes \rho_i, \\ B &:= \sum_{i=1}^n p_i B_i = \sum_{i=1}^n p_i |i\rangle\langle i| \otimes \sigma_i \end{aligned}$$

Note that the operators A_i, B_j , $i, j \in \{1, 2, \dots, n\}$, have orthogonal support for $i \neq j$, i.e.,

$$\text{Tr} A_i A_j = 0 = \text{Tr} A_i B_j = \text{Tr} B_i B_j \quad \text{for } i \neq j. \quad (21)$$

Let π_A and π_i denote the orthogonal projectors onto the support of A and ρ_i , respectively, with $i = 1, 2, \dots, n$. Then using (21) and the convexity of the function $-\log x$, we obtain

$$\begin{aligned} D_{\min}(A||B) &= -\log \text{Tr}(\pi_A B) = -\log \left(\sum_{i=1}^n p_i \text{Tr}(\pi_i \sigma_i) \right) \\ &\leq \sum_{i=1}^n p_i [-\log \text{Tr}(\pi_i \sigma_i)] \\ &= \sum_{i=1}^n p_i D_{\min}(\rho_i||\sigma_i). \end{aligned} \quad (22)$$

Taking the partial traces of A and B over \mathbf{C}^n yields the operators

$$\mathrm{Tr}_{\mathbf{C}^n} A = \sum_i p_i \rho_i \quad ; \quad \mathrm{Tr}_{\mathbf{C}^n} B = \sum_i p_i \sigma_i.$$

However, since the partial trace over \mathbf{C}^n is a CPTP map, we have by Lemma 7 that

$$D_{\min}\left(\sum_{i=1}^n p_i \rho_i \middle| \middle| \sum_{i=1}^n p_i \sigma_i\right) \leq D_{\min}(A||B) \quad (23)$$

The inequalities (22) and (23) yield the joint convexities:

$$D_{\min}\left(\sum_{i=1}^n p_i \rho_i \middle| \middle| \sum_{i=1}^n p_i \sigma_i\right) \leq \sum_{i=1}^n p_i D_{\min}(\rho_i || \sigma_i) \quad (24)$$

■

Lemma 9 *The max- relative entropy of two mixtures of states, $\rho := \sum_{i=1}^n p_i \rho_i$ and $\sigma := \sum_{i=1}^n p_i \sigma_i$, satisfies the following bound:*

$$D_{\max}(\rho || \sigma) \leq \max_{1 \leq i \leq n} D_{\max}(\rho_i || \sigma_i) \quad (25)$$

Proof By definition (8):

$$D_{\max}(\rho || \sigma) = \log \min\{\lambda : \mathrm{Tr}[P_+^\lambda(\rho - \lambda\sigma)] = 0\},$$

where $P_+^\lambda = \{\rho \geq \lambda\sigma\}$. Consider the projection operators $P_+^{\lambda,i} := \{\rho_i \geq \lambda\sigma_i\}$ for $i = 1, 2, \dots, n$. Then

$$\begin{aligned} 0 \leq \mathrm{Tr}[P_+^\lambda(\rho - \lambda\sigma)] &= \sum_i p_i \mathrm{Tr}[P_+^\lambda(\rho_i - \lambda\sigma_i)] \\ &\leq \sum_i p_i \mathrm{Tr}[P_+^{\lambda,i}(\rho_i - \lambda\sigma_i)], \end{aligned} \quad (26)$$

by Lemma 1. Set $\lambda = \max_{1 \leq i \leq n} \lambda_i$ where for each $i = 1, 2, \dots, n$, λ_i is defined by

$$\log \lambda_i = D_{\max}(\rho_i || \sigma_i).$$

For this choice of λ , each term in the sum on the right hand side of (26) vanishes, implying that $\mathrm{Tr}[P_+^\lambda(\rho - \lambda\sigma)] = 0$, and hence $\lambda \geq D_{\max}(\rho || \sigma)$. ■

Lemma 10 *The min- and max- relative entropies of two states ρ and σ are related to the quantum relative entropy $S(\rho || \sigma) := \mathrm{Tr}[\rho \log \rho - \rho \log \sigma]$ as follows:*

$$D_{\min}(\rho || \sigma) \leq S(\rho || \sigma) \leq D_{\max}(\rho || \sigma). \quad (27)$$

Proof We first prove the upper bound $S(\rho || \sigma) \leq D_{\max}(\rho || \sigma)$:

Let $\rho \leq 2^\alpha \sigma$, with $\alpha = D_{\max}(\rho || \sigma)$. Then using the operator monotonicity of the logarithm [29], we have $\log \rho \leq$

$\alpha + \log \sigma$. This in turn implies that $\rho \log \sigma \geq \rho \log \rho - \alpha \rho$. Hence, for a state ρ , $\mathrm{Tr} \rho \log \sigma \geq \mathrm{Tr} \rho \log \rho - \alpha$, and

$$\begin{aligned} S(\rho || \sigma) &:= \mathrm{Tr} \rho \log \rho - \mathrm{Tr} \rho \log \sigma \\ &\leq \mathrm{Tr} \rho \log \rho - \mathrm{Tr} \rho \log \rho + \alpha \\ &= D_{\max}(\rho || \sigma). \end{aligned} \quad (28)$$

We next prove the bound $D_{\min}(\rho || \sigma) \leq S(\rho || \sigma)$: Consider the CPTP map, \mathcal{T} , defined by

$$\mathcal{T}(\omega) = \pi_\rho \omega \pi_\rho + \bar{\pi}_\rho \omega \bar{\pi}_\rho,$$

where ω is any density matrix, π_ρ is the projector onto the support of ρ , and $\bar{\pi}_\rho = I - \pi_\rho$. Note that $\mathcal{T}(\rho) = \rho$.

Due to the monotonicity of $S(\rho || \sigma)$ under CPTP maps, we have

$$\begin{aligned} S(\rho || \sigma) &\geq S(\mathcal{T}(\rho) || \mathcal{T}(\sigma)) \\ &= \mathrm{Tr} \rho \log \rho - \mathrm{Tr} \rho \log (\pi_\rho \sigma \pi_\rho) \\ &= S(\rho || \pi_\rho \sigma \pi_\rho). \end{aligned} \quad (29)$$

Define the normalized state $\tilde{\sigma} := \frac{1}{c} \pi_\rho \sigma \pi_\rho$ where $c = \mathrm{Tr}(\pi_\rho \sigma)$. Then

$$\begin{aligned} S(\rho || \pi_\rho \sigma \pi_\rho) &= S(\rho || c \tilde{\sigma}) \\ &= \mathrm{Tr} \rho (\log \rho - \log \tilde{\sigma}) - (\log c) \cdot \mathrm{Tr} \rho \\ &= S(\rho || \tilde{\sigma}) - \log c \\ &\geq -\log c = D_{\min}(\rho || \sigma). \end{aligned} \quad (30)$$

From (29) and (30) we conclude that $D_{\min}(\rho || \sigma) \leq S(\rho || \sigma)$. ■

The following lemma is obtained easily from the definitions of the min- and max-relative entropies.

Lemma 11 *The min- and max- relative entropies are invariant under joint unitary transformations.*

Lemma 12 *The min- relative entropy of two states ρ and σ for which $\mathrm{supp} \rho \subseteq \mathrm{supp} \sigma$, satisfies the following bounds:*

$$D_{\min}(\rho || \sigma) \leq D_{\max}(\rho || \sigma) \leq -\log \mu_{\min}(\sigma), \quad (31)$$

where $\mu_{\min}(\sigma)$ denotes the minimum non-zero eigenvalue of σ . Further,

$$D_{\min}(\rho || \sigma) \leq -\log \left[1 - \frac{1}{2} \|\rho - \sigma\|_1\right] \quad (32)$$

Proof The first inequality in (31) has been proved in Lemma 5. Since $\mathrm{supp} \rho \subseteq \mathrm{supp} \sigma$, we have $\pi_\rho \leq \pi_\sigma$, where π_ρ and π_σ denote the projectors onto the supports of ρ and σ respectively. Further, using the bounds $\rho \leq \pi_\rho$ and $\pi_\sigma \leq \mu_{\min}(\sigma)^{-1} \sigma$, where $\mu_{\min}(\sigma)$ denotes the minimum non-zero eigenvalue of σ , we obtain

$$\rho \leq \mu_{\min}(\sigma)^{-1} \sigma.$$

Using the definition of $D_{\max}(\rho||\sigma)$ we therefore infer that $D_{\max}(\rho||\sigma) \leq -\log \mu_{\min}(\sigma)$.

The bound (32) follows from the fact that

$$\text{Tr}(\pi_\rho \sigma) \geq 1 - \frac{1}{2} \|\rho - \sigma\|_1, \quad (33)$$

which can be seen as follows. Consider the scenario of state discrimination. Suppose it is known that a finite quantum system is in one of two states ρ and σ , with equal apriori probability. To determine which state it is in, one does a binary Positive Operator-Valued Measurement (POVM) with elements E_1 and E_2 , and $E_1 + E_2 = I$. If the outcome corresponding to E_1 occurs then the system is inferred to be in the state ρ , whereas if the outcome corresponding to E_2 occurs, then the system is inferred to be in the state σ . The average probability of error in state discrimination is given by

$$\begin{aligned} p_e^{av} &= 1 - \frac{1}{2}(\text{Tr} E_1 \rho + \text{Tr} E_2 \sigma) \\ &= \frac{1}{2}(\text{Tr} E_2 \rho + \text{Tr} E_1 \sigma) \end{aligned} \quad (34)$$

Note that the two terms in the parenthesis, in the last line of (34) are, respectively, the *Type I error* and the *Type II error*, in the language of hypothesis testing. By Helstrom's Theorem [30] the minimum possible value of p_e^{av} is given by

$$\frac{1}{2} [1 - \frac{1}{2} \|\rho - \sigma\|_1].$$

Now consider a POVM in which $E_1 = \pi_\rho$ (the projector onto the support of ρ) and $E_2 = I - \pi_\rho$. In this case, the *Type I error* vanishes and $p_e^{av} = \frac{1}{2} \text{Tr}(\pi_\rho \sigma)$, which by Helstrom's theorem satisfies the bound:

$$\frac{1}{2} \text{Tr}(\pi_\rho \sigma) \geq \frac{1}{2} [1 - \frac{1}{2} \|\rho - \sigma\|_1],$$

hence yielding the desired bound (33). Note that $D_{\min}(\rho||\sigma)$ is therefore related to the average probability of error of state discrimination (between the states ρ and σ) when the Type I error vanishes.

Note that (33) can also be proved algebraically by a simple use of Lemma 1. ■

It is known that if one has asymptotically many copies of two states ρ and σ , the error in discriminating between them decreases exponentially, and the error exponent is given by the so-called quantum Chernoff bound $\xi(\rho, \sigma)$ [22]. The latter has been shown to be given by $\xi(\rho, \sigma) = -\log(\min_{0 \leq s \leq 1} \text{Tr} \rho^s \sigma^{1-s})$. From the expression (10) it follows that the min-relative entropy provides a lower bound to the quantum Chernoff bound, i.e.,

$$\xi(\rho, \sigma) \geq D_{\min}(\rho||\sigma).$$

V. A NEW ENTANGLEMENT MONOTONE

Here we introduce a new entanglement monotone for a bipartite state ρ . We denote it by $E_{\max}(\rho)$ and call it the *max-relative entropy of entanglement*.

In [31], Vedral and Plenio proved a set of sufficient conditions under which a measure $D(\rho||\sigma)$ of the “distance” between two bipartite quantum states, ρ and σ , defines an entanglement monotone, $E(\rho)$, through the following expression:

$$E(\rho) := \min_{\sigma \in \mathcal{D}} D(\rho||\sigma).$$

Here the minimum is taken over the set \mathcal{D} of all separable states (see also [32]). The conditions ensure that (i) $E(\rho) = 0$ if and only if ρ is separable; (ii) $E(\rho)$ is unchanged by a local change of basis and (iii) $E(\rho)$ does not increase on average under local operations and classical communication (LOCC) [34]. By proving that these conditions are satisfied by $D_{\max}(\rho||\sigma)$, we are able to define a new entanglement monotone.

Theorem 1 *For a bipartite state ρ , the quantity*

$$E_{\max}(\rho) := \min_{\sigma \in \mathcal{D}} D_{\max}(\rho||\sigma), \quad (35)$$

where the minimum is taken over the set \mathcal{D} of all separable states, is an entanglement monotone.

Proof It suffices to prove that the max-relative entropy $D_{\max}(\rho||\sigma)$ satisfies the following properties, which constitute the set of sufficient conditions proved in [31].

- $D_{\max}(\rho||\sigma) \geq 0$ with equality if and only if $\rho = \sigma$.
- $D_{\max}(\rho||\sigma) = D_{\max}(U\rho U^\dagger||U\sigma U^\dagger)$ for any unitary operator U .
- $D_{\max}(\text{Tr}_p \rho||\text{Tr}_p \sigma) \leq D_{\max}(\rho||\sigma)$, where Tr_p denotes a partial trace.
- $\sum_i \alpha_i D_{\max}(\rho_i/\alpha_i||\sigma_i/\beta_i) \leq \sum_i D_{\max}(\rho_i||\sigma_i)$, where $\alpha_i = \text{Tr} \rho_i$, $\beta_i = \text{Tr} \sigma_i$, $\rho_i = V_i \rho V_i^\dagger$, $\sigma_i = V_i \sigma V_i^\dagger$ and $\sum_i V_i^\dagger V_i = 1$.
- For any set $\{P_i\}$ of mutually orthogonal projectors, i.e., $P_i P_j = \delta_{ij} P_i$,

$$\begin{aligned} &D_{\max}\left(\sum_i P_i \rho P_i \middle| \middle| \sum_i P_i \sigma P_i\right) \\ &= \sum_i D_{\max}(P_i \rho P_i || P_i \sigma P_i). \end{aligned} \quad (36)$$

- For any projector P ,

$$D_{\max}(\rho \otimes P || \sigma \otimes P) = D_{\max}(\rho||\sigma).$$

From Lemma 6 we have that if $\rho = \sigma$ then $D_{\max}(\rho||\sigma) = 0$. To prove the converse, i.e., $D_{\max}(\rho||\sigma) = 0$ implies that $\rho = \sigma$, note that $D_{\max}(\rho||\sigma) = 0 \implies (\sigma - \rho) \geq 0$. On the other hand, since ρ and σ are states, we have $\text{Tr}(\sigma - \rho) = 0$, which in turn implies that $(\sigma - \rho) = 0$. This completes the proof of the first property. The second property follows from Lemma 11. The third property

follows from Lemma 7, since the partial trace is a CPTP map.

The fourth property can be proved as follows. Using the definition (7) of the max-relative entropy, we have

$$\begin{aligned}
& \sum_i \alpha_i D_{\max}(\rho_i/\alpha_i || \sigma_i/\beta_i) \\
&= \sum_i \alpha_i \log \left[\mu_{\max} \left(\left(\frac{\sigma_i}{\beta_i} \right)^{-\frac{1}{2}} \frac{\rho_i}{\alpha_i} \left(\frac{\sigma_i}{\beta_i} \right)^{-\frac{1}{2}} \right) \right] \\
&= \sum_i \alpha_i \log \left(\frac{\beta_i}{\alpha_i} \mu_{\max} \left(\sigma_i^{-\frac{1}{2}} \rho_i \sigma_i^{-\frac{1}{2}} \right) \right) \\
&= \sum_i \alpha_i \log \frac{\beta_i}{\alpha_i} + \sum_i \alpha_i \log \mu_{\max} \left(\sigma_i^{-\frac{1}{2}} \rho_i \sigma_i^{-\frac{1}{2}} \right) \\
&\leq - \sum_i \alpha_i \log \frac{\alpha_i}{\beta_i} + \sum_i \log \mu_{\max} \left(\sigma_i^{-\frac{1}{2}} \rho_i \sigma_i^{-\frac{1}{2}} \right) \\
&\leq \sum_i \log \mu_{\max} \left(\sigma_i^{-\frac{1}{2}} \rho_i \sigma_i^{-\frac{1}{2}} \right) \\
&= \sum_i D_{\max}(\rho_i || \sigma_i). \tag{37}
\end{aligned}$$

In the above we have made use of the fact that $\{\alpha_i\}$ and $\{\beta_i\}$ are probability distributions, and hence $\sum_i \alpha_i \log \frac{\alpha_i}{\beta_i} \geq 0$ [33]. The fifth and sixth properties are seen to hold by inspection. ■

Lemma 13 *The max-relative entropy of entanglement, E_{\max} , is an upper bound to the relative entropy of entanglement [31]*

$$E(\rho) := \min_{\sigma \in \mathcal{D}} S(\rho || \sigma).$$

Proof Let σ^* be a separable state such that

$$\begin{aligned}
D_{\max}(\rho || \sigma^*) &= \min_{\sigma \in \mathcal{D}} D(\rho || \sigma) \\
&= E_{\max}(\rho)
\end{aligned} \tag{38}$$

By Lemma 10 we have

$$\begin{aligned}
D_{\max}(\rho || \sigma^*) &\geq S(\rho || \sigma^*) \\
&\geq \min_{\sigma \in \mathcal{D}} S(\rho || \sigma) \\
&= E(\rho).
\end{aligned} \tag{39}$$

Properties of $E_{\max}(\rho)$ will be investigated in a forthcoming paper.

VI. SMOOTH MIN- AND MAX- RELATIVE ENTROPIES

Smooth min- and max- relative entropies are generalizations of the above-mentioned relative entropy measures, involving an additional *smoothness* parameter $\varepsilon \geq 0$. For $\varepsilon = 0$, they reduce to the *non-smooth* quantities

Definition 4 *For any $\varepsilon \geq 0$, the ε -smooth min- and max-relative entropies of a bipartite state ρ relative to a state σ are defined by*

$$D_{\min}^{\varepsilon}(\rho || \sigma) := \sup_{\bar{\rho} \in B^{\varepsilon}(\rho)} D_{\min}(\bar{\rho} || \sigma)$$

and

$$D_{\max}^{\varepsilon}(\rho || \sigma) := \inf_{\bar{\rho} \in B^{\varepsilon}(\rho)} D_{\max}(\bar{\rho} || \sigma)$$

where $B^{\varepsilon}(\rho) := \{\bar{\rho} \geq 0 : \|\bar{\rho} - \rho\|_1 \leq \varepsilon, \text{Tr}(\bar{\rho}) \leq \text{Tr}(\rho)\}$.

The following two lemmas are used in the proof of Theorem 2 given below.

Lemma 14 *Let ρ_{AB} and σ_{AB} be density operators, let Δ_{AB} be a positive operator, and let $\lambda \in \mathbb{R}$ such that*

$$\rho_{AB} \leq 2^{\lambda} \cdot \sigma_{AB} + \Delta_{AB}.$$

Then $D_{\max}^{\varepsilon}(\rho_{AB} || \sigma_{AB}) \leq \lambda$ for any $\varepsilon \geq \sqrt{8\text{Tr}(\Delta_{AB})}$.

Lemma 15 *Let ρ_{AB} and σ_{AB} be density operators. Then*

$$D_{\max}^{\varepsilon}(\rho_{AB} || \sigma_{AB}) \leq \lambda$$

for any $\lambda \in \mathbb{R}$ and

$$\varepsilon = \sqrt{8\text{Tr}[\{\rho_{AB} > 2^{\lambda}\sigma_{AB}\}\rho_{AB}]}.$$

The proofs of these lemmas are analogous to the proofs of Lemmas 5 and 6 of [20] and are given in the Appendix for completeness.

VII. SPECTRAL DIVERGENCE RATES

In the quantum information spectrum approach one defines spectral divergence rates, defined below, which can also be viewed as generalizations of the quantum relative entropy.

Definition 5 *Given a sequence of states $\hat{\rho} = \{\rho_n\}_{n=1}^{\infty}$ and a sequence of positive operators $\hat{\sigma} = \{\sigma_n\}_{n=1}^{\infty}$, the quantum spectral sup-(inf-)divergence rates are defined in terms of the difference operators $\Pi_n(\gamma) = \rho_n - 2^{n\gamma}\sigma_n$ as*

$$\overline{D}(\hat{\rho} || \hat{\sigma}) := \inf \left\{ \gamma : \limsup_{n \rightarrow \infty} \text{Tr}[\{\Pi_n(\gamma) \geq 0\}\Pi_n(\gamma)] = 0 \right\} \tag{40}$$

$$\underline{D}(\hat{\rho} || \hat{\sigma}) := \sup \left\{ \gamma : \liminf_{n \rightarrow \infty} \text{Tr}[\{\Pi_n(\gamma) \geq 0\}\Pi_n(\gamma)] = 1 \right\} \tag{41}$$

respectively.

Although the use of sequences of states allows for immense freedom in choosing them, there remain a number of basic properties of the quantum spectral divergence rates that hold for all sequences. These are stated and proved in [4]. In the i.i.d. case the sequence is generated from product states $\rho = \{\rho^{\otimes n}\}_{n=1}^{\infty}$, which is used to relate the spectral entropy rates for the sequence ρ to the entropy of a single state ρ .

Note that the above definitions of the spectral divergence rates differ slightly from those originally given in (38) and (39) of [13]. However, they are equivalent, as stated in the following two propositions [35]. For their proofs see [4] or [20].

Proposition 1 *The spectral sup-divergence rate $\overline{D}(\hat{\rho}||\hat{\sigma})$ is equal to*

$$\overline{D}(\hat{\rho}||\hat{\sigma}) := \inf \left\{ \alpha : \limsup_{n \rightarrow \infty} \text{Tr}[\{\rho_n \geq 2^{n\alpha} \sigma_n\} \rho_n] = 0 \right\} \quad (42)$$

which is the original definition of the spectral sup-divergence rate. Hence the two definitions are equivalent.

Proposition 2 *The spectral inf-divergence rate $\underline{D}(\hat{\rho}||\hat{\sigma})$ is equivalent to*

$$\underline{D}(\hat{\rho}||\hat{\sigma}) = \sup \left\{ \alpha : \liminf_{n \rightarrow \infty} \text{Tr}[\{\rho_n \geq 2^{n\alpha} \sigma_n\} \rho_n] = 1 \right\} \quad (43)$$

which is the original definition of the spectral inf-divergence rate. Hence the two definitions are equivalent.

Despite these equivalences, it is useful to use the definitions (40) and (41) for the divergence rates as they allow the application of Lemmas 1, 2 and 3 in deriving various properties of these rates.

The spectral generalizations of the von Neumann entropy, the conditional entropy and the mutual information can all be expressed as spectral divergence rates with appropriate substitutions for the sequence of operators $\hat{\sigma} = \{\sigma_n\}_{n=1}^{\infty}$.

A. Definition of spectral entropy rates

Consider a sequence of Hilbert spaces $\{\mathcal{H}_n\}_{n=1}^{\infty}$, with $\mathcal{H}_n = \mathcal{H}^{\otimes n}$. For any sequence of states $\hat{\rho} = \{\rho_n\}_{n=1}^{\infty}$, with ρ_n being a density matrix acting in the Hilbert space \mathcal{H}_n , the sup- and inf- spectral entropy rates are defined as follows:

$$\begin{aligned} \overline{S}(\hat{\rho}) &= \inf \left\{ \gamma : \liminf_{n \rightarrow \infty} \text{Tr}[\{\rho_n \geq 2^{-n\gamma} I_n\} \rho_n] = 1 \right\} \quad (44) \\ \underline{S}(\hat{\rho}) &= \sup \left\{ \gamma : \limsup_{n \rightarrow \infty} \text{Tr}[\{\rho_n \geq 2^{-n\gamma} I_n\} \rho_n] = 0 \right\}. \end{aligned} \quad (45)$$

Here I_n denotes the identity operator acting in \mathcal{H}_n . These are obtainable from the spectral divergence rates as follows [see [4]]:

$$\overline{S}(\hat{\rho}) = -\underline{D}(\hat{\rho}||\hat{I}); \quad \underline{S}(\hat{\rho}) = -\overline{D}(\hat{\rho}||\hat{I}), \quad (46)$$

where $\hat{I} = \{I_n\}_{n=1}^{\infty}$ is a sequence of identity operators.

It is known that the spectral entropy rates of $\hat{\rho}$ are related to the von Neumann entropies of the states ρ_n as follows:

$$\underline{S}(\hat{\rho}) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} S(\rho_n) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} S(\rho_n) \leq \overline{S}(\hat{\rho}). \quad (47)$$

Moreover, for a sequence of product states $\hat{\rho} = \{\rho^{\otimes n}\}_{n=1}^{\infty}$:

$$\underline{S}(\hat{\rho}) = \lim_{n \rightarrow \infty} \frac{1}{n} S(\rho_n) = \overline{S}(\hat{\rho}). \quad (48)$$

For sequences of bipartite states $\hat{\rho} = \{\rho_n^{AB}\}_{n=1}^{\infty}$, with $\rho_n^{AB} \in \mathcal{B}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n})$, the conditional spectral entropy rates are defined as follows:

$$\overline{S}(A|B) := -\underline{D}(\hat{\rho}^{AB}||\hat{I}^A \otimes \hat{\rho}^B); \quad (49)$$

$$\underline{S}(A|B) := -\overline{D}(\hat{\rho}^{AB}||\hat{I}^A \otimes \hat{\rho}^B). \quad (50)$$

In the above, $\hat{I}^A = \{I_n^A\}_{n=1}^{\infty}$ and $\hat{\rho}^A = \{\rho_n^A\}_{n=1}^{\infty}$, with I_n^A being the identity operator acting in $\mathcal{H}_A^{\otimes n}$ and $\rho_n^A = \text{Tr}_B \rho_n^{AB}$, the partial trace being taken on the Hilbert space $\mathcal{H}_B^{\otimes n}$.

Similarly, the mutual information rates are given by

$$\overline{S}(A : B) := \overline{D}(\hat{\rho}^{AB}||\hat{\rho}^A \otimes \hat{\rho}^B); \quad (51)$$

$$\underline{S}(A|B) := \underline{D}(\hat{\rho}^{AB}||\hat{\rho}^A \otimes \hat{\rho}^B). \quad (52)$$

These spectral entropy rates have several interesting properties (see e.g.[4]) and also have the operational significance of being related to the optimal rates of protocols (see the discussion in the Introduction and the references quoted there).

VIII. RELATION BETWEEN SPECTRAL DIVERGENCE RATES AND SMOOTH MIN- AND MAX- RELATIVE ENTROPIES

In this section we prove the relations between the spectral divergence rates and the smooth relative entropies. As mentioned in the Introduction, the proofs are entirely self-contained, relying only on the definitions of the entropic quantities involved, and the lemmas stated in Section II.

A. Relation between $\overline{D}(\hat{\rho}||\hat{\sigma})$ and the smooth max-relative entropy

Theorem 2 *Given a sequence of bipartite states $\hat{\rho} = \{\rho_n\}_{n=1}^{\infty}$, and a sequence of positive operators $\hat{\sigma} = \{\sigma_n\}_{n=1}^{\infty}$, where $\rho_n, \sigma_n \in \mathcal{B}(\mathcal{H}^{\otimes n})$, the sup-spectral divergence rate $\overline{D}(\hat{\rho}||\hat{\sigma})$, defined by (40) (or equivalently by (42)), satisfies*

$$\overline{D}(\hat{\rho}||\hat{\sigma}) = \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} D_{\max}^{\epsilon}(\rho_n||\sigma_n), \quad (53)$$

where $D_{\max}^\varepsilon(\rho_n||\sigma_n)$ is the smooth max-entropy of the state ρ_n of the sequence $\hat{\rho}$, and the operator σ_n of the sequence $\hat{\sigma}$.

Proof

We first prove the bound

$$\overline{D}(\hat{\rho}||\hat{\sigma}) \geq \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} D_{\max}^\varepsilon(\rho_n||\sigma_n), \quad (54)$$

Let $\delta > 0$ be arbitrary but fixed, and define

$$\gamma := \overline{D}(\hat{\rho}||\hat{\sigma}) + \delta. \quad (55)$$

Then from Proposition 1 it follows that

$$\limsup_{n \rightarrow \infty} \text{Tr}[\{\rho_n^{AB} \geq 2^{n\gamma} \sigma_n^{AB}\} \rho_n^{AB}] = 0. \quad (56)$$

In particular, for any $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\begin{aligned} \text{Tr}[\{\rho_n^{AB} > 2^{n\gamma} \sigma_n^{AB}\} \rho_n^{AB}] &\leq \text{Tr}[\{\rho_n^{AB} \geq 2^{n\gamma} \sigma_n^{AB}\} \rho_n^{AB}] \\ &< \frac{\varepsilon^2}{8}. \end{aligned} \quad (57)$$

Using Lemma 15 we then infer that for all $n \geq n_0$

$$D_{\max}^\varepsilon(\rho_n^{AB}||\sigma_n^{AB}) \leq n\gamma \quad (58)$$

and, hence

$$\limsup_{n \rightarrow \infty} \frac{1}{n} D_{\max}^\varepsilon(\rho_n^{AB}||\sigma_n^{AB}) \leq \gamma. \quad (59)$$

Since this holds for any $\varepsilon > 0$, we conclude

$$\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} D_{\max}^\varepsilon(\rho_n^{AB}||\sigma_n^{AB}) \leq \gamma = \overline{D}(\hat{\rho}||\hat{\sigma}) + \delta. \quad (60)$$

The assertion (54) then follows because this holds for any arbitrary $\delta > 0$.

We next prove the bound

$$\overline{D}(\hat{\rho}||\hat{\sigma}) \leq \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} D_{\max}^\varepsilon(\rho_n||\sigma_n), \quad (61)$$

Let $\tilde{\rho}_{n,\varepsilon}$ be the operator for which

$$D_{\max}(\tilde{\rho}_{n,\varepsilon}||\sigma_n) = \inf_{\tilde{\rho} \in B^\varepsilon(\rho_n)} D_{\max}(\tilde{\rho}||\rho_n) = D_{\max}^\varepsilon(\rho_n||\sigma_n). \quad (62)$$

This implies, in particular, that for any λ for which $\log \lambda \geq D_{\max}^\varepsilon(\rho_n||\sigma_n)$, we have

$$\text{Tr}[\{\tilde{\rho}_{n,\varepsilon} \geq \lambda \sigma_n\}(\tilde{\rho}_{n,\varepsilon} - \lambda \sigma_n)] = 0. \quad (63)$$

For any real constant $\gamma > 0$, let us define the projection operator

$$P_n^\gamma := \{\rho_n \geq 2^{n\gamma} \sigma_n\}. \quad (64)$$

Note that

$$\begin{aligned} \text{Tr}[P_n^\gamma \rho_n] &= \text{Tr}[P_n^\gamma \tilde{\rho}_{n,\varepsilon}] + \text{Tr}[P_n^\gamma (\rho_n - \tilde{\rho}_{n,\varepsilon})] \\ &\leq \text{Tr}[P_n^\gamma (\tilde{\rho}_{n,\varepsilon} - 2^{n\alpha} \sigma_n)] + 2^{n\alpha} \text{Tr}[P_n^\gamma \sigma_n] \\ &\quad + \text{Tr}[\{\rho_n \geq \tilde{\rho}_{n,\varepsilon}\}(\rho_n - \tilde{\rho}_{n,\varepsilon})] \\ &\leq \text{Tr}[\{\tilde{\rho}_{n,\varepsilon} \geq 2^{n\alpha} \sigma_n\}(\tilde{\rho}_{n,\varepsilon} - 2^{n\alpha} \sigma_n)] \\ &\quad + 2^{n(\alpha-\gamma)} + \varepsilon \end{aligned} \quad (65)$$

In the above we have made use of Lemma 1, Lemma 2 and Corollary 1.

Let $\lambda := 2^{n\alpha}$ and choose $\log \lambda = D_{\max}^\varepsilon(\rho_n||\sigma_n) + \delta/2$ for any arbitrary $\delta > 0$. Further let us choose $\gamma = \alpha + \delta/2$. Then, by (63), the first term on the right hand side of (65) vanishes. Moreover, the second term also goes to zero as $n \rightarrow \infty$. Therefore, for n large enough and any $\delta' > 0$, in the limit $\varepsilon \rightarrow 0$, we must have that

$$\text{Tr}(P_n^\gamma \rho_n) \leq \delta'. \quad (66)$$

This together with Proposition 1 implies that $\gamma \geq \overline{D}(\hat{\rho}||\hat{\sigma})$. The required bound (61) then follows from the choice of the parameters α and γ . ■

B. Relation between $\underline{D}(\hat{\rho}||\hat{\sigma})$ and the smooth min-relative entropy

Theorem 3 Given a sequence of bipartite states $\hat{\rho} = \{\rho_n\}_{n=1}^\infty$, and a sequence of positive operators $\hat{\sigma} = \{\sigma_n\}_{n=1}^\infty$, where $\rho_n, \sigma_n \in \mathcal{B}(\mathcal{H}^{\otimes n})$, the inf-spectral divergence rate $\underline{D}(\hat{\rho}||\hat{\sigma})$, defined by (41) (or equivalently by (43)), satisfies

$$\underline{D}(\hat{\rho}||\hat{\sigma}) = \lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} D_{\min}^\varepsilon(\rho_n||\sigma_n), \quad (67)$$

where $D_{\min}^\varepsilon(\rho_n||\sigma_n)$ is the smooth min-relative entropy of the state ρ_n of the sequence $\hat{\rho}$ and the operator σ_n of the sequence $\hat{\sigma}$.

Proof From the definition (43) of $\underline{D}(\hat{\rho}||\hat{\sigma})$ it follows that for any $\gamma \leq \underline{D}(\hat{\rho}||\hat{\sigma})$ and any $\delta > 0$, for n large enough

$$\text{Tr}[P_n^\gamma \rho_n] > 1 - \delta, \quad (68)$$

where $P_n^\gamma := \{\rho_n \geq 2^{n\gamma} \sigma_n\}$.

For any given $\alpha > 0$, choose $\gamma = \underline{D}(\hat{\rho}||\hat{\sigma}) - \alpha$, and let

$$\tilde{\rho}_{n,\gamma} := P_n^\gamma \rho_n P_n^\gamma \quad (69)$$

Then using (68) and the ‘‘Gentle measurement lemma’’, Lemma 4, we infer that, for n large enough, $\tilde{\rho}_{n,\gamma} \in B^\varepsilon(\rho_n)$ with $\varepsilon = 2\sqrt{\delta}$. Let $\pi_{n,\gamma}$ denote the projection onto the support of $\tilde{\rho}_{n,\gamma}$.

We first prove bound

$$\lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} D_{\min}^\varepsilon(\rho_n||\sigma_n) \geq \underline{D}(\hat{\rho}||\hat{\sigma}). \quad (70)$$

For n large enough,

$$\begin{aligned} D_{\min}^{\varepsilon}(\rho_n || \sigma_n) &= \sup_{\tilde{\rho}_n \in B^{\varepsilon}(\rho_n)} D_{\min}(\tilde{\rho}_n || \sigma_n) \\ &\geq D_{\min}(\tilde{\rho}_{n,\gamma} || \sigma_n) \\ &= -\log \text{Tr}(\pi_{n,\gamma} \sigma_n) \\ &\geq -\log \text{Tr}(P_{n,\gamma} \sigma_n) \geq n\gamma. \end{aligned} \quad (71)$$

The last inequality in (71) follows from Lemma 2. Hence, for n large enough,

$$\frac{1}{n} D_{\min}^{\varepsilon}(\rho_n || \sigma_n) \geq \gamma = \underline{D}(\hat{\rho} || \hat{\sigma}) - \alpha, \quad (72)$$

and since α is arbitrary, we obtain the desired bound (70).

To complete the proof of Theorem 3, we assume that

$$\lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} D_{\min}^{\varepsilon}(\rho_n || \sigma_n) > \underline{D}(\hat{\rho} || \hat{\sigma}), \quad (73)$$

and prove that this leads to a contradiction.

Let $\tilde{\rho}_{n,\varepsilon}$ be the operator for which

$$D_{\min}^{\varepsilon}(\rho_n || \sigma_n) = D_{\min}(\tilde{\rho}_{n,\varepsilon} || \sigma_n) = -\log \text{Tr}(\tilde{\pi}_{n,\varepsilon} \sigma_n), \quad (74)$$

where $\tilde{\pi}_{n,\varepsilon}$ is the projection onto the support of $\tilde{\rho}_{n,\varepsilon}$.

Note that

$$\begin{aligned} \text{Tr}(\tilde{\pi}_{n,\varepsilon} \rho_n) &= \text{Tr}[\tilde{\pi}_{n,\varepsilon}((\rho_n - \tilde{\rho}_{n,\varepsilon}) + \tilde{\rho}_{n,\varepsilon})] \\ &= \text{Tr}[\tilde{\pi}_{n,\varepsilon}(\rho_n - \tilde{\rho}_{n,\varepsilon})] + \text{Tr}\tilde{\rho}_{n,\varepsilon} \\ &\geq \text{Tr}[\{\rho_n \leq \tilde{\rho}_{n,\varepsilon}\}(\rho_n - \tilde{\rho}_{n,\varepsilon})] + \text{Tr}[\tilde{\rho}_{n,\varepsilon}] \\ &\geq -\varepsilon + 1 - \varepsilon = 1 - 2\varepsilon. \end{aligned} \quad (75)$$

We arrive at the second last line of (75) using Lemma 1. The last inequality is obtained by using the fact that $\tilde{\rho}_{n,\varepsilon} \in B^{\varepsilon}(\rho_n)$, and the bound

$$\text{Tr}[\{\rho_n \leq \tilde{\rho}_{n,\varepsilon}\}(\rho_n - \tilde{\rho}_{n,\varepsilon})] \geq -\varepsilon,$$

which arises from the fact that $\tilde{\rho}_{n,\varepsilon} \in B^{\varepsilon}(\rho_n)$.

Define,

$$\begin{aligned} \beta_{\varepsilon} &:= \liminf_{n \rightarrow \infty} \left[\frac{1}{n} D_{\min}(\tilde{\rho}_{n,\varepsilon} || \sigma_n) \right] \\ &:= \liminf_{n \rightarrow \infty} \left[-\frac{1}{n} \log \text{Tr}(\tilde{\pi}_{n,\varepsilon} \sigma_n) \right] \end{aligned} \quad (76)$$

and

$$\gamma := \lim_{\varepsilon \rightarrow 0} \beta_{\varepsilon}.$$

Obviously, $\beta_{\varepsilon} \geq \gamma$. Note that the assumption (73) is equivalent to the assumption $\gamma > \underline{D}(\hat{\rho} || \hat{\sigma})$. Let γ_0 be such that

$$\beta_{\varepsilon} > \gamma_0 > \underline{D}(\hat{\rho} || \hat{\sigma}). \quad (77)$$

Then, for any fixed $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$

$$\frac{1}{n} D_{\min}(\tilde{\rho}_{n,\varepsilon} || \sigma_n) \geq \beta_{\varepsilon}. \quad (78)$$

The above inequality can be rewritten as

$$\text{Tr}(\tilde{\pi}_{n,\varepsilon} \sigma_n) \leq 2^{-n\beta_{\varepsilon}} \quad (79)$$

Using (79) we obtain the following:

$$\begin{aligned} \text{Tr}(\tilde{\pi}_{n,\varepsilon} \rho_n) &= \text{Tr}[\tilde{\pi}_{n,\varepsilon}(\rho_n - 2^{n\gamma_0} \sigma_n)] + 2^{n\gamma_0} \text{Tr}[\tilde{\pi}_{n,\varepsilon} \sigma_n] \\ &\leq \text{Tr}[\{\rho_n \geq 2^{n\gamma_0} \sigma_n\}(\rho_n - 2^{n\gamma_0} \sigma_n)] + 2^{n\gamma_0} 2^{-n\beta_{\varepsilon}} \\ &\leq \text{Tr}[\{\rho_n \geq 2^{n\gamma_0} \sigma_n\}(\rho_n - 2^{n\gamma_0} \sigma_n)] + 2^{-n(\beta_{\varepsilon} - \gamma_0)} \end{aligned} \quad (80)$$

The second term on the right hand side of (80) tends to zero asymptotically in n , since $\beta_{\varepsilon} > 0$. However, the first term does not tend to 1, since $\gamma_0 > \underline{D}(\hat{\rho} || \hat{\sigma})$ by assumption. Hence we obtain the bound

$$\text{Tr}(\tilde{\pi}_{n,\varepsilon} \rho_n) < 1 - c_0, \quad (81)$$

for some constant $c_0 > 0$. This contradicts (75) in the limit $\varepsilon \rightarrow 0$. \blacksquare

IX. APPENDIX

Proof of Lemma 14

Proof Define

$$\begin{aligned} \alpha_{AB} &:= 2^{\lambda} \cdot \sigma_{AB} \\ \beta_{AB} &:= 2^{\lambda} \cdot \sigma_{AB} + \Delta_{AB}. \end{aligned}$$

and

$$T_{AB} := \alpha_{AB}^{\frac{1}{2}} \beta_{AB}^{-\frac{1}{2}}.$$

Let $|\Psi\rangle = |\Psi\rangle_{ABR}$ be a purification of ρ_{AB} and let $|\Psi'\rangle := T_{AB} \otimes I_R |\Psi\rangle$ and $\rho'_{AB} := \text{Tr}_R(|\Psi'\rangle\langle\Psi'|)$.

Note that

$$\begin{aligned} \rho'_{AB} &= T_{AB} \rho_{AB} T_{AB}^{\dagger} \\ &\leq T_{AB} \beta_{AB} T_{AB}^{\dagger} \\ &= \alpha_{AB} = 2^{\lambda} \cdot \sigma_{AB}, \end{aligned}$$

which implies $D_{\max}(\rho'_{AB} || \sigma_{AB}) \leq \lambda$. It thus remains to be shown that

$$\|\rho_{AB} - \rho'_{AB}\|_1 \leq \sqrt{8\text{Tr}(\Delta_{AB})}. \quad (82)$$

We first show that the Hermitian operator

$$\bar{T}_{AB} := \frac{1}{2}(T_{AB} + T_{AB}^{\dagger}).$$

satisfies

$$\bar{T}_{AB} \leq I_{AB}. \quad (83)$$

For any vector $|\phi\rangle = |\phi\rangle_{AB}$,

$$\begin{aligned}\|T_{AB}|\phi\rangle\|^2 &= \langle\phi|T_{AB}^\dagger T_{AB}|\phi\rangle = \langle\phi|\beta_{AB}^{-\frac{1}{2}}\alpha_{AB}\beta_{AB}^{-\frac{1}{2}}|\phi\rangle \\ &\leq \langle\phi|\beta_{AB}^{-\frac{1}{2}}\beta_{AB}\beta_{AB}^{-\frac{1}{2}}|\phi\rangle = \|\phi\|^2\end{aligned}$$

where the inequality follows from $\alpha_{AB} \leq \beta_{AB}$. Hence, for any vector $|\phi\rangle$,

$$\begin{aligned}\|\bar{T}_{AB}|\phi\rangle\| &\leq \frac{1}{2}\|T_{AB}|\phi\rangle + T_{AB}^\dagger|\phi\rangle\| \\ &\leq \frac{1}{2}\|T_{AB}|\phi\rangle\| + \frac{1}{2}\|T_{AB}^\dagger|\phi\rangle\| \leq \|\phi\|,\end{aligned}$$

which implies (83).

We now determine the overlap between $|\Psi\rangle$ and $|\Psi'\rangle$,

$$\begin{aligned}\langle\Psi|\Psi'\rangle &= \langle\Psi|T_{AB} \otimes I_R|\Psi\rangle \\ &= \text{Tr}(|\Psi\rangle\langle\Psi|T_{AB} \otimes I_R) = \text{Tr}(\rho_{AB}T_{AB}).\end{aligned}$$

Because ρ_{AB} has trace one, we have

$$\begin{aligned}1 - |\langle\Psi|\Psi'\rangle| &\leq 1 - \Re\langle\Psi|\Psi'\rangle = \text{Tr}(\rho_{AB}(I_{AB} - \bar{T}_{AB})) \\ &\leq \text{Tr}(\beta_{AB}(I_{AB} - \bar{T}_{AB})) \\ &= \text{Tr}(\beta_{AB}) - \text{Tr}(\alpha_{AB}^{\frac{1}{2}}\beta_{AB}^{\frac{1}{2}}) \\ &\leq \text{Tr}(\beta_{AB}) - \text{Tr}(\alpha_{AB}) = \text{Tr}(\Delta_{AB}).\end{aligned}$$

Here, the second inequality follows from the fact that, because of (83), the operator $I_{AB} - \bar{T}_{AB}$ is positive, and $\rho_{AB} \leq \beta_{AB}$. The last inequality holds because $\alpha_{AB}^{\frac{1}{2}} \leq \beta_{AB}^{\frac{1}{2}}$, which is a consequence of the operator monotonicity of the square root (Proposition V.1.8 of [29]).

Using (5) and the fact that the fidelity between two pure states is given by their overlap, we find

$$\begin{aligned}\| |\Psi\rangle\langle\Psi| - |\Psi'\rangle\langle\Psi'| \|_1 &\leq 2\sqrt{2(1 - |\langle\Psi|\Psi'\rangle|)} \\ &\leq 2\sqrt{2\text{Tr}(\Delta_{AB})} \leq \varepsilon.\end{aligned}$$

Inequality (82) then follows because the trace distance can only decrease when taking the partial trace. ■

Proof of Lemma 15

Proof Let Δ_{AB}^+ and Δ_{AB}^- be mutually orthogonal positive operators such that

$$\Delta_{AB}^+ - \Delta_{AB}^- = \rho_{AB} - 2^\lambda \sigma_{AB}.$$

Furthermore, let P_{AB} be the projector onto the support of Δ_{AB}^+ , i.e.,

$$P_{AB} = \{\rho_{AB} > 2^\lambda \sigma_{AB}\}.$$

We then have

$$\begin{aligned}P_{AB}\rho_{AB}P_{AB} &= P_{AB}(\Delta_{AB}^+ + 2^\lambda \sigma_{AB} - \Delta_{AB}^-)P_{AB} \\ &\geq \Delta_{AB}^+\end{aligned}$$

and, hence,

$$\sqrt{8\text{Tr}(\Delta_{AB}^+)} \leq \sqrt{8\text{Tr}(P_{AB}\rho_{AB})} = \varepsilon.$$

The assertion now follows from Lemma 14 because

$$\rho_{AB} \leq 2^\lambda \sigma_{AB} + \Delta_{AB}^+.$$

■

Acknowledgements

The author is very grateful to Milan Mosonyi for carefully reading the paper and pointing out an error in the first version. She would also like to thank Reinhard Werner and Tony Dorlas for helpful suggestions.

-
- [1] R. Renner, “Security of quantum key distribution,” PhD thesis, ETH Zurich, *arXiv:quant-ph/0512258*, 2005.
 - [2] S. Kullback and R. Leibler, “On information and sufficiency,” *Ann. Math. Stat.*, vol. 22, pp. 79-86, 1951.
 - [3] H. Umegaki, “Conditional expectations in an operator algebra IV (entropy and information),” *Kodai Math. Sem. Rep.*, vol. 14, pp. 59-85.
 - [4] G. Bowen and N. Datta, “Beyond i.i.d. in quantum information theory,” *arXiv:quant-ph/0604013*, *Proceedings of the 2006 IEEE International Symposium on Information Theory*, 2006.
 - [5] T. S. Han, *Information-Spectrum Methods in Information Theory*, Springer-Verlag, 2002.
 - [6] T. S. Han and S. Verdú, “Approximation theory of output statistics,” *IEEE Trans. Inform. Theory*, vol. 39, pp. 752-772, 1993.
 - [7] H. Nagaoka and M. Hayashi, “An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses,” *arXiv:quant-ph/0206185*, 2002.
 - [8] T. Ogawa and H. Nagaoka, “Strong converse and Stein’s lemma in quantum hypothesis testing,” *IEEE Trans. Inform. Theory*, vol. 46, pp. 2428-2433, 2000.
 - [9] S. Verdú and T. S. Han, “A general formula for channel capacity,” *IEEE Trans. Inform. Theory*, vol. 40, pp. 1147-1157, 1994.
 - [10] G. Bowen and N. Datta, “Quantum coding theorems for arbitrary sources, channels and entanglement resources,” *arXiv:quant-ph/0610003*, 2006.
 - [11] G. Bowen and N. Datta, “Asymptotic entanglement manipulation of bipartite pure states,” *IEEE Trans. Inform. Theory*, vol. 54, pp. 3677-3686, 2008.
 - [12] G. Bowen and N. Datta, “Entanglement cost for sequences of arbitrary quantum states,” *arXiv:0704.1957*, 2007.

- [13] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1753–1768, 2003.
- [14] M. Hayashi, "General formulas for fixed-length quantum entanglement concentration," *IEEE Trans. Inform. Theory*, Vol. 52, No. 5, 1904-1921, 2006.
- [15] K. Matsumoto, "Entanglement cost and distillable entanglement of symmetric states," *arXiv:0708.3129*, 2007.
- [16] R. Renner and R. Koenig, "Universally composable privacy amplification against quantum adversaries," *Proc. of TCC 2005, LNCS, Springer*, vol. 3378, 2005.
- [17] R. Renner and S. Wolf, "Smooth Rényi entropy and applications," in *Proc. International Symposium on Information Theory*, p. 233. IEEE, 2004.
- [18] R. Renner and S. Wolf, "Simple and tight bounds for information reconciliation and privacy amplification," in *Advances in Cryptology — ASIACRYPT 2005*, LNCS, vol. 3788, pp. 199–216. Springer, 2005.
- [19] R. Renner, S. Wolf, and J. Wullschleger, "Trade-offs in information-theoretic multi-party one-way key agreement," to appear in *Proc. of ICITS 2007*, 2007.
- [20] N. Datta and R. Renner, "Smooth Rényi Entropies and the Quantum Information Spectrum," *arXiv:0801.0282*; to appear in *IEEE Trans. Inf. Theory*, 2009.
- [21] M. Ohya and D. Petz, "Quantum Entropy and Its Use," Springer-Verlag, 1993.
- [22] K.M.R. Audenaert et al. "The Quantum Chernoff Bound," *Phys. Rev. Letts.*, vol. 98, 160501, 2007.
- [23] R. Koenig, R. Renner and C. Schaffner, "The operational meaning of min- and max- entropy," *arXiv:0807.1338*, 2008.
- [24] M. Mosonyi and N. Datta, "Generalized relative entropies and the capacities of classical-quantum channels," *arXiv:0810.3478*, 2008.
- [25] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.
- [26] T. Ogawa and H. Nagaoka, "New proof of the channel coding theorem via hypothesis testing in quantum information theory," *arXiv:quant-ph/0208139*, 2002.
- [27] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Inf. Theory*, vol. 45, 1999.
- [28] M. Hayashi, "Quantum Information An Introduction," Springer, 2006.
- [29] R. Bhatia, *Matrix Analysis*, Springer.
- [30] C.W. Helstrom, *Quantum Detection and Estimation Theory*. Mathematics in Science and Engineering, vol. 123, Academic Press, London 1976.
- [31] V. Vedral and M.B. Plenio, "Entanglement measures and purification procedures," *Phys. Rev. A*, vol. 57, pp. 1147–1157, 1998.
- [32] M. Horodecki, M. Horodecki, M. Horodecki, J. Oppenheim, A. Sen(De), U. Sen and B. Synak, "Local versus non-local information in quantum information theory: formalism and phenomenon," *Phys. Rev. A*, Vol. 71, No. 5, 062325, 2005.
- [33] T.M. Cover and J.A. Thomas, "Elements of Information Theory," *A Wiley-Interscience Publication*, 1991.
- [34] If Λ denotes a LOCC map and $\{p_i, \rho_i\}$ denotes an ensemble of states such that ρ_i is obtained with probability p_i under the action of Λ on ρ , then $E(\rho) \geq \sum_i p_i E(\rho_i)$.
- [35] Note that in [4] and [13], the logarithm was taken to base e , whereas here we take the logarithm to base 2.

Biography

Nilanjana Datta received a Ph.D. degree from ETH Zurich, Switzerland, in 1996. From 1997 to 2000, she was a postdoctoral researcher at the Dublin Institute of Advanced Studies, C.N.R.S. Marseille, and EPFL in Lausanne. In 2001 she joined the University of Cambridge, as a Lecturer in Mathematics of Pembroke College, and a member of the Statistical Laboratory, in the Centre for Mathematical Sciences. She is currently an Affiliated Lecturer of the Faculty of Mathematics, University of Cambridge, and a Fellow of Pembroke College.